

Leveraging beginning programming to introduce data science

MC²: Data Science across the CSU

Alex Chen

CSU Dominguez Hills

February 12, 2020

Outline

- My background, our department, our students
- MAT 241 Programming and Technology for Teaching Secondary School Mathematics
- Future directions for the data science concentration at CSUDH

My background

- Went to grad school at UCLA, working in image processing with Andrea Bertozzi.
- Moved to a postdoc at UNC Chapel Hill, working on mathematical biology.
- Worked for 4 years at GE Global Research, on finance, computer vision, and additive manufacturing.
- In my second year at CSUDH.
- I'm trying to bring an industry perspective as much as possible to my teaching.



CSUDH and the math department

- Historically have trained many math teachers.
- Once had an M.A. in Mathematics Teaching
- Current: about 1/3 math option, 2/3 math education option majors, about 160 majors, with a growing trend toward the math option.
- Future: Transition to introducing a data science concentration option.



Plan to introduce the data science option

- Recently hired 3.5 faculty who work in applied math or statistics.
- Advantages:
 - Interest from students: defense, actuarial science, business
 - Great math club
 - Strong programs at CSUDH in biology, computer science, psychology, business
 - Math department politics
 - Investments toward science and technology
- Challenges:
 - Low awareness of new data science classes.
 - Not enough data science classes.
 - Not much room in existing curriculum for electives.

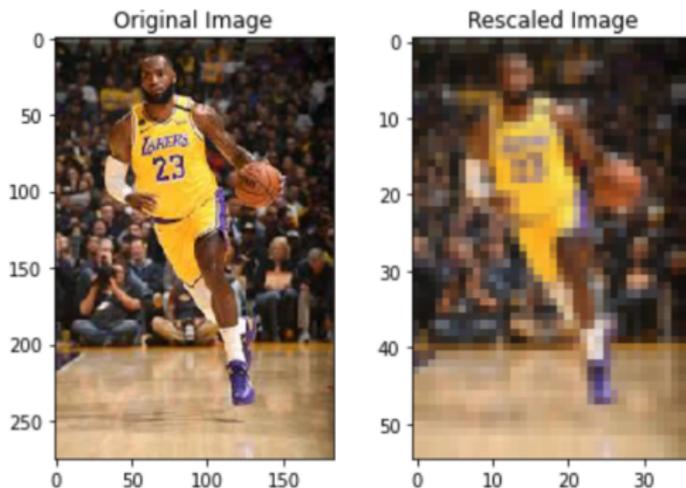


MAT 241, introductory programming

- MAT 241 Programming and Technology for Teaching Secondary School Mathematics is an existing course for math education concentration majors.
- Prerequisite is Calculus II.
- Has traditionally focused on different tools for teachers:
 - Scratch (intro programming language for kids)
 - Spreadsheets
 - LaTeX
 - Python
 - Small introduction to other tools like R or Matlab
- A bit disjointed, all parties thought not enough time on basic programming.
- Required for math education concentration, not for math concentration.
- A big challenge is the differing programming background of the (35) students.

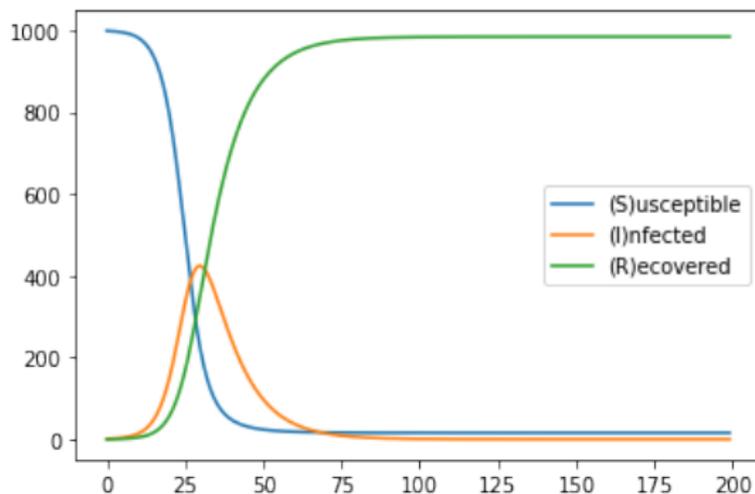
Transitioning MAT 241 to a new form

- May be able to serve as a pilot for introduction to data science.
- Focus $\frac{1}{2}$ to $\frac{2}{3}$ of the class on basic programming using Python.
- Remainder on math modeling with Python, consisting of different topics of interest in data science.
- Examples:
 - Signal and image processing
 - Epidemiology
 - Elections



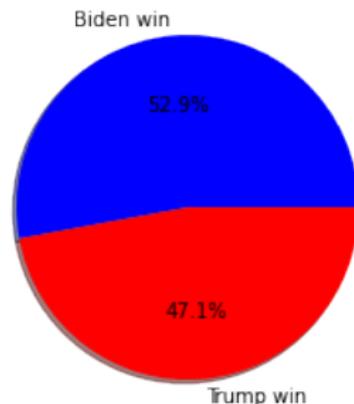
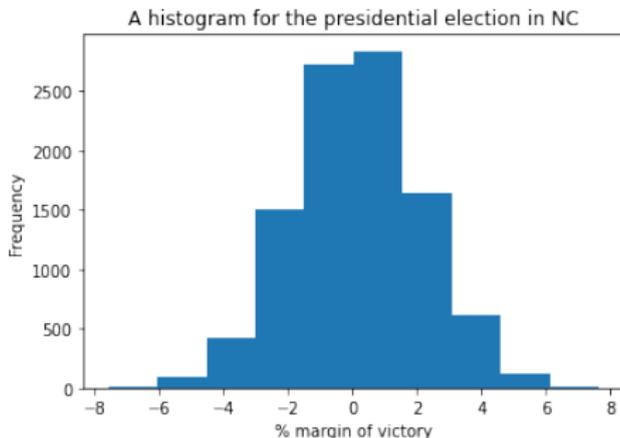
Examples of math modeling modules with Python

- Epidemiological modeling with SIR models
- Students are introduced to the basic differential equations.
- Code is mostly done for them; they change parameters and make observations.



Examples of math modeling modules with Python

- Election simulation with state-by-state simulations.
- Simulations using random normal distributions and polling data.



```
[ ] # On trial 0, the electoral votes for each (270 required to win):  
    print("Biden got %d electoral votes" % Biden_votes[0])  
    print("Trump got %d electoral votes" % Trump_votes[0])
```

```
Biden got 351 electoral votes  
Trump got 187 electoral votes
```

Future directions for the data science option

- Use MAT 241 as a recruiting tool to get students interested into the data science option.
- Better integration with solving problems for other classes?
- Introduce more classes in data science and allow them as options for both math education and math concentrations.
- Better communication to students about career options and class options.

Questions? Comments? Recommendations?
achen at csudh dot edu

